



Published in final edited form as:

*Am J Prev Med.* 2015 June ; 48(6): 747–754. doi:10.1016/j.amepre.2014.12.010.

## Evaluating Diabetes Health Policies Using Natural Experiments:

### The Natural Experiments for Translation in Diabetes Study

**Ronald T. Ackermann, MD, MPH, O. Kenrik Duru, MD, MS, Jeanine B. Albu, MD, Julie A. Schmittiel, PhD, Stephen B. Soumerai, ScD, James F. Wharam, MB, BCh, BAO, MPH, Mohammed K. Ali, MBChB, MSc, MBA, Carol M. Mangione, MD, MSPH, Edward W. Gregg, PhD, and on behalf of the NEXT-D Study Group**

Departments of Medicine and Medical Social Sciences (Ackermann), Northwestern University Feinberg School of Medicine, Chicago, Illinois; Division of General Internal Medicine/Health Services Research (Duru, Mangione), University of California, Los Angeles; Division of Research (Schmittiel), Kaiser Permanente Northern California, Oakland, California; Department of Medicine and the New York Obesity Research Center (Albu), St. Luke's-Roosevelt Hospital Center, New York, New York; Department of Population Medicine (Soumerai, Wharam), Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts; Hubert Department of Global Health (Ali), Rollins School of Public Health, Emory University; and Division of Diabetes Translation (Ali, Gregg), National Center for Chronic Disease Prevention and Health Promotion, CDC, Atlanta, Georgia

### Abstract

The high prevalence and costs of type 2 diabetes makes it a rapidly evolving focus of policy action. Health systems, employers, community organizations, and public agencies have increasingly looked to translate the benefits of promising research interventions into innovative policies intended to prevent or control diabetes. Though guided by research, these health policies provide no guarantee of effectiveness and may have opportunity costs or unintended consequences. Natural experiments use pragmatic and available data sources to compare specific policies to other policy alternatives or predictions of what would likely have happened in the absence of any intervention. The Natural Experiments for Translation in Diabetes (NEXT-D) Study is a network of academic, community, industry, and policy partners, collaborating to advance the methods and practice of natural experimental research, with a shared aim of identifying and prioritizing the best policies to prevent and control diabetes. This manuscript describes the NEXT-D Study group's multi-sector natural experiments in areas of diabetes prevention or control as case examples to illustrate the selection, design, analysis, and challenges inherent to natural experimental study approaches to inform development or evaluation of health policies.

---

Address correspondence to: Ronald T. Ackermann, MD, MPH, Northwestern University Feinberg School of Medicine, 750 North Lake Shore Drive, 10th Floor, Chicago IL 60611. r.ackermann@northwestern.edu.

No financial disclosures were reported by the authors of this paper.

## Introduction

About 29 million Americans have diabetes, and another 86 million have prediabetes, placing them at high risk for progression to type 2 diabetes as well as additional health problems, lower quality of life, and higher medical and non-medical costs.<sup>1</sup> Recent increases in the burden of type 2 diabetes have encouraged rapid policy action by both public agencies and private organizations. Optimistically, many agencies and organizations have attempted to implement interventions derived from seminal research studies, often making adaptations to enhance access, feasibility, and scalability. For example, recent policy action established a national recognition program at CDC to define standards for how community organizations should offer lifestyle-based diabetes prevention programs consistent with the Diabetes Prevention Program (DPP) clinical trial.<sup>2</sup> Similarly, value-based health insurance designs are being implemented by some health systems to reduce or eliminate out-of-pocket costs for preventive services or medications that have strong evidence of benefit.<sup>3</sup>

Although the use of past research to inform diabetes-related health policies is encouraging, there is no guarantee that an intervention derived from a successful clinical trial will be effective or cost effective under real-life circumstances. For example, clinical trials have demonstrated that intensive interventions to support diabetes self-management improve glycemic control and cardiovascular risk factors,<sup>4,5</sup> but policies by health payers to fund disease management interventions in diabetes and many other areas have yielded mixed results.<sup>6,7</sup> Similarly, duplicative prescribing and poly-pharmacy can increase adverse medication events, but policies that cap the number of medications covered by state-based public assistance programs have been linked to lower use of recommended glucose-lowering agents, increased hospitalization, and higher public expenditures.<sup>8,9</sup> Unintended policy consequences, such as the potential to exacerbate rather than reduce health disparities, are also an important target for health policy research.<sup>10,11</sup> Ideally, when a health policy is enacted, ongoing strong research would establish whether the benefits outweigh harms, and if costs are reasonable for society.<sup>12</sup> Unfortunately, many health policies still remain unevaluated or are assessed using research designs that are vulnerable to considerable bias and are thus potentially misleading.<sup>13</sup>

## Selecting Valid Research Designs for a Policy Intervention

Although randomized trials are the gold standard for clinical efficacy research, they are often impractical or inappropriate to test policy effectiveness.<sup>14,15</sup> Conversely, natural experiments use pragmatic research designs and readily available data sources to evaluate and compare a new or existing policy to other policy alternatives or predictions of what may have happened in the absence of any intervention.<sup>14,16,17</sup> Natural experiments have been applied to numerous disciplines, including political science, education, psychology, economics, history, and sociology.<sup>17–19</sup> Within the context of health research, natural experiments have been particularly influential in policies related to indoor tobacco laws, Medicaid drug reimbursement policies, and provision of vouchers for health and social services.<sup>20–24</sup>

For each policy intervention, there is typically more than one choice for a natural experimental study design, with the goal of implementing the strongest design feasible given

the study question, implementation context, and data sources available. Table 1 provides examples of several common design options, including how adequately they address common threats to validity.

Individually randomized trials are often difficult to implement but should still be considered in the evaluation of public or organizational policies.<sup>24,25</sup> When individual randomization is not feasible, adaptations to the RCT design, such as cluster-randomized parallel-group or stepped-wedge intervention trials can provide strong alternatives by allocating groups to interventions at a provider, clinical practice, organizational, community, or regional level.<sup>26,27</sup> Importantly, such designs must carefully consider decreases in the effective sample size, the possibility of imbalance among treatment groups when randomized clusters are small, and the need for special statistical models to account for correlation of outcomes within clusters.<sup>27</sup>

If no form of randomization is feasible or appropriate, the interrupted time series design (with and without control series), and to a lesser extent, pre–post studies with a strongly matched comparison group provide alternatives that can control for most common threats to validity.<sup>17</sup> These designs must carefully consider common threats such as bias and confounding and the role of complementary statistical approaches to address these threats during the analysis. Examples of such approaches include multivariable adjustment or propensity matching with or without adjustment. Another important strategy to reduce uncertainty for a particular study design is to replicate findings by repeating the analysis in another context (i.e., external validation) or through sensitivity analyses that examine whether similar results are achieved using alternative comparison groups or study design options. Although used in many past policy evaluations, simple cross-sectional or pre–post designs without a comparison group are generally considered weak for evaluating intervention effectiveness. As reflected in Table 1, these designs are attractive because of their simplicity but can be misleading, as they are unable to control for common biases such as history and selection.

## The Natural Experiments for Translation in Diabetes Study

The Natural Experiments for Translation in Diabetes (NEXT-D) study is a dynamic collaboration among five academic centers, numerous organizational and policy partners, CDC, and the National Institute of Digestive and Diabetes and Kidney Diseases. NEXT-D centers include the Harvard Pilgrim Health Care Institute and Harvard Medical School, Kaiser Permanente Northern California, Northwestern University, New York St. Luke's-Roosevelt Hospital, and the University of California, Los Angeles (UCLA). Each of these centers has an evaluation partnership with one or more non-academic entities to evaluate a public or organizational policy area using a natural experimental study design that is practical, methodologically rigorous, and suited to the needs and interests of each policy stakeholder. By including policy decision makers in the network, NEXT-D seeks to address the most relevant gaps in the scientific research literature while developing data collection, analysis, interpretation, and communication strategies to improve policy decision making. Partnering directly with policy stakeholders also increases the likelihood that research findings will be translated into real-life policies and practice.<sup>28</sup>

The NEXT-D study group adopted a conceptual framework to portray the relationships among NEXT-D projects. This framework was derived from existing social-ecologic heuristics<sup>29,30</sup> and considers how each natural experiment might be affected by other interventions or contextual influences across geographic, organizational, and economic sectors. Using the framework, the study group considers how best to design natural experiments to evaluate how policies are adopted and implemented, whom they reach, and whether they are effective and sustainable. The goal for each natural experiment is a rigorously designed quasi-experimental study that carefully considers the individual and synergistic effects of influences at multiple policy levels.<sup>15,16</sup> These interactive effects are depicted in Figure 1 and explored briefly in the text that follows.

A central focus of the framework is to understand the intended or unintended impacts of policies enacted at any level. These impacts can be related to diabetes incidence, progression, or health and economic outcomes, as well as the distribution of those outcomes (e.g., reducing or increasing health disparities) throughout the population. Policies impact these key outcomes either directly or indirectly through effects that ultimately must culminate in successful changes in the behaviors of individuals or others such as healthcare providers, public health professionals, community health workers, or peers.

The health system is portrayed as the smallest sphere, underscoring its modest yet important influence on overall population health outcomes relative to other sectors, primarily via behavioral, preventive, and health-care services. Examples include clinical practice recommendations, public reporting of quality metrics, new resources for health promotion and self-management support, and the implementation of electronic health record systems to coordinate healthcare and behavioral support services.<sup>31,32</sup> At a slightly broader level, the sponsors or purchasers of healthcare services, such as public agencies (e.g., Centers for Medicare and Medicaid Services [CMS], state Medicaid offices) and employers, are depicted as investing in efforts (implemented either directly or via health plan partners) to improve patient and provider behaviors, and to offer additional resources and services to support those behaviors. Examples include the design of new coverage benefits, special programs to support prevention and care management, and restructuring payments to incentivize behaviors that lead to better health and health care.<sup>33–36</sup>

Individuals and populations also regularly interact with community resources from religious, civic, socially driven, or commercial entities. Community resources are a crucial influence on health-related behaviors and health status. For example, faith-based organizations play an increasing role in social support, self-management, and even clinical services for people with diabetes or prediabetes.<sup>37</sup>

Finally, the public policy environment may play a strong role in influencing health through actions such as the development of laws and regulations to shape the economic, physical, social, and cultural environments in which people live, work, and recreate. Public policies may result in both intended and unintended consequences through direct regulatory impacts or via interactions with other interventions that are deployed simultaneously across other levels.

## Examples of Existing Natural Experiments Supported by NEXT-D

Diverse study designs are employed in NEXT-D to determine the impacts of policies enacted at multiple levels. Below, we describe very briefly the existing NEXT-D projects as case examples for the selection, design, analysis, and challenges of using robust natural experimental research designs to inform the evolution of policies in public health and healthcare contexts.

### Example 1: Kaiser Permanente Northern California

Telephonic coaching to encourage healthy lifestyle choices is a population-based approach to wellness and diabetes prevention that is being explored by many health plans and purchasers. As gestational diabetes mellitus (GDM) has increased in incidence, more health plans are also implementing strategies to encourage glucose screening for earlier identification and treatment of both type 2 diabetes and prediabetes postpartum. Kaiser Permanente Northern California (KPNC) has partnered with a large healthcare purchaser to assess the impact of providing population-based telephonic wellness coaching, as well as a postpartum diabetes risk-reduction intervention targeting women who have had GDM.<sup>38</sup> These studies use multiple data sources and interrupted time series with control series, as well as pre-post longitudinal control group designs to assess patient satisfaction, perceptions of the programs, BMI, smoking-cessation rates, hemoglobin A1c and fasting plasma glucose levels, and onset of diabetes in populations. The studies will also examine potential impacts on health care and health disparities, as well as the efficiency and programmatic costs to help stakeholders understand the barriers and facilitators to expanding these programs within the health plan if they are found to be effective.

### Example #2: University of California, Los Angeles

At UCLA, investigators are partnering with the Innovations Group at United Healthcare, the largest commercial health insurance company in the U.S., to evaluate the health and economic impacts of a new diabetes-specific health insurance benefits package. The Diabetes Health Plan (DHP), which is now being offered by several dozen large- and medium-sized employers throughout the U.S., providing chronic condition management, financial incentives, and improved access to value-based services (e.g., established preventive medications, treatments, and services at very low or no cost) for patients with prediabetes and diabetes.<sup>39</sup> The study uses a longitudinal quasi-experimental design with a propensity-matched comparison group of employers who do not offer the DHP to evaluate the effect on adherence to evidence-based medications, use of emergency department or hospital inpatient services, rates of progression from prediabetes to diabetes, and total costs of care. UCLA meets regularly with United's Innovations Group to ensure that the study's findings will help United make decisions about further DHP refinements and its replication. This study also will determine what employers should expect when implementing this type of disease-specific plan in real-world settings, where program uptake and employee engagement can vary.

**Example #3: St. Luke's-Roosevelt Health System**

Investigators at St. Luke's-Roosevelt are collaborating with six primary care clinics in New York City to study pragmatic interventions at the health system/provider level that address challenges to early diabetes detection and prevention in primary care practice. Interventions use chronic care model principles and involve electronic health record modifications and integrated health care systems/community linkages that are rolled out across each clinic in a stepped fashion, allowing for comparisons over time as well as across implementing and non-implementing clinics. Using electronic health record data sources and a time series study design, this study will compare rates of test completion and detection of diabetes or prediabetes. A pre-post with propensity-matched comparison group study design will be used to determine the effects of the enhanced management interventions on clinical intermediate outcomes (e.g., hemoglobin A1c), for about 1,500 adults who are identified as having prediabetes or type 2 diabetes who receive care at a practice location implementing (versus not implementing) the enhanced interventions.<sup>40</sup> Evaluation results will inform the continuing implementation and wider replication of enhanced management intervention strategies to other clinics in the network.

**Example #4: Harvard Pilgrim Health Care Institute**

More than 40% of workers now have high-deductible health plans (HDHPs),<sup>41</sup> and recent policies are expected to spur further increases because of coverage mandates, incentives for plans with lower actuarial value (non-“Cadillac” plans), and greater upfront affordability.<sup>42</sup> It is not yet known whether HDHPs will have their intended effect of creating “activated healthcare consumers” who change their behaviors and seek higher-value services to lower their own health expenditures or, rather, who simply reduce utilization, even for care that is considered essential. Harvard partnered with OptumInsight Life Sciences, Inc., to develop a national data set involving a 13-year rolling sample (2001–2013) of 1.3 million adults with diabetes from all 50 states who are health plan members of an HDHP or more traditional plans (controls). The study is employing an interrupted time series with comparison series study design to assess the impact of HDHPs on diabetes control, risk factors, preventive care, and health outcomes such as emergency department visits and preventable hospitalizations; utilization, including outpatient visits; and costs.<sup>43–45</sup> Results of this work could be used to design health plan features that promote more equitable high-quality care and better outcomes among patients with diabetes.

**Example #5: Northwestern University Feinberg School of Medicine**

To address the growing incidence of type 2 diabetes in the U.S., UnitedHealth Group (UHG), Y-USA, and CDC have partnered to bring a group-based adaptation of the DPP lifestyle intervention to a national scale. This initiative combines actions at the purchaser, community resource, and public policy levels and leverages efforts by CDC and Y-USA to build the capacity of community organizations nationwide to deliver a standardized lifestyle intervention program derived from the DPP.<sup>46</sup> These interventions ultimately target several million eligible adults with prediabetes in almost 50 large geographic markets throughout the U.S. The study will use an interrupted time series with propensity-matched comparison series design to determine the relative effects of UHG outreach/engagement activities and



local efforts by YMCAs to enroll participants within different demographic subgroups. By integrating additional data from Y-USA, the project will examine how implementation varies across regions, and using national UHG claims data sources, effectiveness will be evaluated through changes in prescription expenditures; obesity-related risk factors such as total cholesterol and blood glucose; healthcare utilization; and total healthcare expenditures.<sup>46</sup>

## Discussion

New public and organizational policies are being enacted every day to improve health and behavior, but we continually miss important opportunities to evaluate those policies using valid research designs. Natural experiments are forms of pragmatic research that use readily available data sources to evaluate and compare the intended and unintended outcomes of new or existing policies, rather than simply assume they are a wise use of limited resources. No study design is without limitation; thus, the process of selecting a natural experimental approach must carefully consider the trade-offs between feasibility and scientific validity.

Ideally, the design of a natural experiment should be deliberated within the context of a partnership that includes capable evaluators, policy decision-makers, and other stakeholders. Together, these partners must consider important choices and ultimately select a design that will be feasible to implement while providing organizational leaders and other individuals with valid and actionable outcomes, such as who can be reached by new initiatives; whether all target groups have equitable access; if benefits vary across different vulnerable population subgroups; if there are unintended consequences; and the amount and types of expenditures that are needed to implement those initiatives “at scale.” Such research can be considered one badly needed form of implementation science, which seeks to elucidate how best to replicate and scale evidence-based programs across diverse populations and differing settings. Partners may also collaborate to devise creative strategies for overcoming common pragmatic research challenges, such as the paucity of information available in administrative data systems to determine the characteristics of target populations (e.g., race/ethnicity or SES). Overcoming these challenges might require linkages with other data sources, such as neighborhood-level data regarding socioeconomic indicators, or the use of individual surname analysis done prior to de-identifying the data for analysis. These are but a few of the examples of the common challenges that can be solved collaboratively through research partnerships, or even networks of partnerships, as is the case in NEXT-D.

With an eye on quantifying the intended and unintended impacts of population-targeted diabetes prevention and control initiatives, the NEXT-D network is positioned to provide empirical and pragmatic information to guide the ongoing evolution of policies and practices to improve diabetes care and prevention. Through the expanded use of strong natural experimental methods, NEXT-D evaluates public health and health-care policies while those policies are being implemented “unperturbed,” in a fashion that truly reflects health impacts for real-world populations and at natural levels of “exposure” (or “dosage”). Although the individual natural experimental methods adopted by NEXT-D sites are not entirely novel, the support of a network of collaborating academic, health system, public health, commercial, and community stakeholders to advance the evolution of health policies in chronic disease care and prevention is quite unique. In addition, although NEXT-D focuses on diabetes, its

collaborative approach for conceptualizing the policy environment, engaging policy partners in the design and application of natural experiments, and transferring findings back to key stake-holders in ways that promote continual progress toward more efficient and effective population health initiatives is not only relevant but of paramount importance to the entire policy community.

## Acknowledgments

The Natural Experiments for Translation in Diabetes (NEXT-D) Study is funded jointly by CDC and the NIH/ National Institute of Diabetes and Digestive and Kidney Diseases under grant numbers U58-DP002717, DP002718, DP002719, DP002721, and DP002722. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of CDC or the NIH.

The NEXT-D Study Group: Harvard Medical School and Harvard Pilgrim Healthcare Institute: Frank Wharam, MB, BCh, BAO MPH (Co-Principal Investigator [PI]), Steve Soumerai, ScD, MSPH, Connie Trinacty, PhD, Emma Eggleston, MD, MPH, Fang Zhang, PhD, Christine Lu, PhD, Robert LeCates, MA, Claire Canning, MA, Dennis Ross-Degnan, ScD, MSPH (Co-PI); Kaiser Permanente Northern California: Julie A. Schmittiel, PhD (PI), Assiamira Ferrara, MD, PhD, Romain S. Neugebauer, PhD, Sara Adams; Northwestern University: Ronald T. Ackermann, MD, MPH (PI), Joyce Tang, MD, Dustin D. French, PhD, Andrew Cooper, MS, Raymond Kang, MS, David T. Liss, PhD, Margaret Moran, MPH, and Indiana University: Chandan K. Saha, PhD, Ann M. Holmes, PhD; St. Luke's-Roosevelt: Jeanine B. Albu, MD (PI), Nancy Sohler, PhD, Brenda Matti, MD, Edwin Young, MD, Gary Burke, MD, F.X. Pi-Sunyer, MD, Nicholas Freudenberg, PhD, Jordan Sill, MS, Dan Baxter, MD; University of California, Los Angeles: Carol M. Mangione, MD, MSPH (Co-PI), O. Kenrik Duru, MD, MS (Co-PI), Susan Ettner, PhD, Lindsay Kimbro, MPP, Norman Turk, MS, Jinnan Li, MPH, Ekatarina Vaisberg, Neil Steers, PhD; CDC Division of Diabetes Translation: Edward W. Gregg, PhD, Mohammed K. Ali, MBChB, MSc, MBA (also Emory University), Bernice Moore, MBA, Meda E. Pavkov, MD, Rui Li, PhD, Theodore Thompson, MS, Ping Zhang, PhD; National Institute of Diabetes and Digestive and Kidney Diseases: Sanford A. Garfield, PhD, Christine Hunter, PhD.

The NEXT-D Study Group would also like to acknowledge the support of our partners and collaborators, including UnitedHealthcare (Anyia Kirvan, RN, MS, Bob Luchs, MS, MPH, Abigail Keckhafer, MBA, MPH, Charles Chan, CMA, MS); Y-USA (Matt M. Longjohn, MD MPH, Rachel L. Karabenick, Heather Hodge, Katherine H. Hohman); United-Health Center for Health Reform and Modernization (Deneen Vojta, MD, Tom Beauregard); the Diabetes Prevention and Control Alliance at UnitedHealth Group (Joel French, Satish Bandapati); OptumInsight Life Sciences, Inc.; additional partners at Kaiser Permanente Northern California (Vicki George, Andi Fargeix, Nancy Goler MD, Rashel Sanna, Mindy Boccio, David J. Bellamy); and healthcare providers and executive leadership affiliated with St. Luke's-Roosevelt and partnering delivery systems (particularly Sylvaine Frances, PA, CDE, Eva Johansson RN, VP, SLRHC, City University of New York students, the William F. Ryan Community Health Center [Barbara Hood and Dustin Prybzilla], Alex Botta, PhD, Julian Botta, BS, and Fera Ps Inc.).

## References

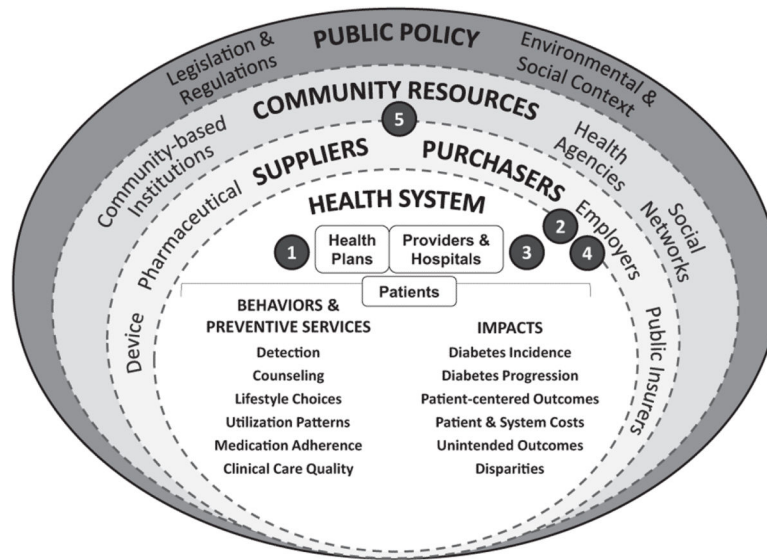
1. CDC.. National diabetes statistical report: estimates of diabetes and its burden in the United States, 2014. DHHS; CDC; Atlanta, GA: 2014. <http://www.cdc.gov/diabetes/pubs/statsreport14/national-diabetes-report-web.pdf>
2. CDC.. National Diabetes Prevention Program. CDC; Atlanta, GA: 2012. [www.cdc.gov/diabetes/prevention/index.htm](http://www.cdc.gov/diabetes/prevention/index.htm)
3. Fendrick AM. Value-based insurance design for diabetes mellitus: approaches to optimal pharmacoeconomic implementation. *Am J Manag Care*. 2010; 16((11) (suppl)):S314–S322. [PubMed: 21517647]
4. Deakin T, McShane CE, Cade JE, Williams RD. Group based training for self-management strategies in people with type 2 diabetes mellitus. *Cochrane Database Syst Rev*. 2005; (2):CD003417. [PubMed: 15846663]
5. Norris SL, Lau J, Smith SJ, Schmid CH, Engelgau MM. Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. *Diabetes Care*. 2002; 25(7):1159–1171. <http://dx.doi.org/10.2337/diacare.25.7.1159>. [PubMed: 12087014]



6. Chodosh J, Morton SC, Mojica W, et al. Meta-analysis: chronic disease self-management programs for older adults. *Ann Intern Med.* 2005; 143(6):427–438. <http://dx.doi.org/10.7326/0003-4819-143-6-200509200-00007>. [PubMed: 16172441]
7. Peikes D, Chen A, Schore J, Brown R. Effects of care coordination on hospitalization, quality of care, and health care expenditures among Medicare beneficiaries: 15 randomized trials. *JAMA.* 2009; 301(6):603–618. <http://dx.doi.org/10.1001/jama.2009.126>. [PubMed: 19211468]
8. Bishop CE, Ryan AM, Gilden DM, Kubisiak J, Thomas CP. Effect of an expenditure cap on low-income seniors' drug use and spending in a state pharmacy assistance program. *Health Serv Res.* 2009; 44(3):1010–1028. <http://dx.doi.org/10.1111/j.1475-6773.2009.00951.x>. [PubMed: 19291168]
9. Soumerai S. Unintended outcomes of medicaid drug cost-containment policies on the chronically mentally ill. *J Clin Psychiatry.* 2003; 64(suppl 17):19–22.
10. Woolf SH, Braveman P. Where health disparities begin: the role of social and economic determinants—and why current policies may make matters worse. *Health Aff (Millwood).* 2011; 30(10):1852–1859. <http://dx.doi.org/10.1377/hlthaff.2011.0685>. [PubMed: 21976326]
11. McHugh MD, Carthon JM, Kang XL. Medicare readmissions policies and racial and ethnic health disparities: a cautionary tale. *Policy Polit Nurs Pract.* 2010; 11(4):309–316. <http://dx.doi.org/10.1177/1527154411398490>. [PubMed: 21531966]
12. Wharam JF, Daniels N. Toward evidence-based policy making and standardized assessment of health policy reform. *JAMA.* 2007; 298(6):676–679. <http://dx.doi.org/10.1001/jama.298.6.676>. [PubMed: 17684191]
13. Majumdar SR, Soumerai SB. The unhealthy state of health policy research. *Health Aff (Millwood).* 2009; 28(5):w900–w908. <http://dx.doi.org/10.1377/hlthaff.28.5.w900>. [PubMed: 19671571]
14. Shadish WR, Cook TD. The renaissance of field experimentation in evaluating interventions. *Annu Rev Psychol.* 2009; 60:607–629. <http://dx.doi.org/10.1146/annurev.psych.60.110707.163544>. [PubMed: 18616391]
15. West SG, Duan N, Pequegnat W, et al. Alternatives to the randomized controlled trial. *Am J Public Health.* 2008; 98(8):1359–1366. <http://dx.doi.org/10.2105/AJPH.2007.124446>. [PubMed: 18556609]
16. Craig P, Cooper C, Gunnell D, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health.* 2012; 66(12):1182–1186. <http://dx.doi.org/10.1136/jech-2011-200375>. [PubMed: 22577181]
17. Shadish, WR.; Cook, TD.; Campbell, DT. Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin; Boston: 2001. p. xxip. 623
18. Durlauf, SN.; Blume, L. The New Palgrave Dictionary of Economics. 2nd ed. Palgrave Macmillan; New York: 2008.
19. Petticrew M, Cummins S, Ferrell C, et al. Natural experiments: an underused tool for public health? *Public Health.* 2005; 119(9):751–757. <http://dx.doi.org/10.1016/j.puhe.2004.11.008>. [PubMed: 15913681]
20. Callinan JE, Clarke A, Doherty K, Kelleher C. Legislative smoking bans for reducing secondhand smoke exposure, smoking prevalence and tobacco consumption. *Cochrane Database Syst Rev.* 2010; (4):CD005992. [PubMed: 20393945]
21. Costello EJ, Compton SN, Keeler G, Angold A. Relationships between poverty and psychopathology: a natural experiment. *JAMA.* 2003; 290(15):2023–2029. <http://dx.doi.org/10.1001/jama.290.15.2023>. [PubMed: 14559956]
22. Sargent RP, Shepard RM, Glantz SA. Reduced incidence of admissions for myocardial infarction associated with public smoking ban: before and after study. *BMJ.* 2004; 328(7446):977–980. <http://dx.doi.org/10.1136/bmj.38055.715683.55>. [PubMed: 15066887]
23. Soumerai SB, Avorn J, Ross-Degnan D, Gortmaker S. Payment restrictions for prescription drugs under Medicaid. Effects on therapy, cost, and equity. *N Engl J Med.* 1987; 317(9):550–556. <http://dx.doi.org/10.1056/NEJM198708273170906>. [PubMed: 3302713]
24. Baicker K, Taubman SL, Allen HL, et al. The Oregon experiment—effects of Medicaid on clinical outcomes. *N Engl J Med.* 2013; 368(18):1713–1722. <http://dx.doi.org/10.1056/NEJMs1212321>. [PubMed: 23635051]

25. Ludwig J, Sanbonmatsu L, Gennetian L, et al. Neighborhoods, obesity, and diabetes—a randomized social experiment. *N Engl J Med*. 2011; 365(16):1509–1519. <http://dx.doi.org/10.1056/NEJMsa1103216>. [PubMed: 22010917]
26. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol*. 2006; 6:54. <http://dx.doi.org/10.1186/1471-2288-6-54>. [PubMed: 17092344]
27. Donner, A.; Klar, N. *Design and Analysis of Cluster Randomization Trials in Health Research*. Arnold; Co-published by the Oxford University Press; New York, NY: 2000. p. xp. 178
28. Schmittdiel JA, Grumbach K, Selby JV. System-based participatory research in health care: an approach for sustainable translational research and quality improvement. *Ann Fam Med*. 2010; 8(3):256–259. <http://dx.doi.org/10.1370/afm.1117>. [PubMed: 20458110]
29. Stokols D. Translating social ecological theory into guidelines for community health promotion. *Am J Health Promot*. 1996; 10(4):282–298. <http://dx.doi.org/10.4278/0890-1171-10.4.282>. [PubMed: 10159709]
30. Whittemore R, Melkus GD, Grey M. Applying the social ecological theory to type 2 diabetes prevention and management. *J Community Health Nurs*. 2004; 21(2):87–99. [http://dx.doi.org/10.1207/s15327655jchn2102\\_03](http://dx.doi.org/10.1207/s15327655jchn2102_03). [PubMed: 15123438]
31. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. *JAMA*. 2002; 288(15):1909–1914. <http://dx.doi.org/10.1001/jama.288.15.1909>. [PubMed: 12377092]
32. Mangione CM, Gerzoff RB, Williamson DF, et al. The association between quality of care and the intensity of diabetes disease management programs. *Ann Intern Med*. 2006; 145(2):107–116. <http://dx.doi.org/10.7326/0003-4819-145-2-200607180-00008>. [PubMed: 16847293]
33. Agency for Healthcare Research and Quality. *Designing and Implementing Medicaid Disease and Care Management Programs: A User's Guide*. Rockville, MD: Mar. 2008 2008. Report No.: AHRQ Publication No. 07(08)-0063
34. Brown R, Peikes D, Chen A, Schore J. 15-site randomized trial of coordinated care in Medicare FFS. *Health Care Financ Rev*. 2008; 30(1):5–25.
35. Chernew ME, Rosen AB, Fendrick AM. Value-based insurance design. *Health Aff (Millwood)*. 2007; 26(2):w195–w203. <http://dx.doi.org/10.1377/hlthaff.26.2.w195>. [PubMed: 17264100]
36. Spaulding A, Fendrick AM, Herman WH, et al. A controlled trial of value-based insurance design —The MHealthy: Focus on Diabetes (FOD) trial. *Implement Sci*. 2009; 4:19. <http://dx.doi.org/10.1186/1748-5908-4-19>. [PubMed: 19351413]
37. Newlin K, Dyess SM, Allard E, Chase S, Gail D'Eramo M. A methodological review of faith-based health promotion literature: advancing the science to expand delivery of diabetes education to black Americans. *J Relig Health*. 2011; 51(4):1075–1097. <http://dx.doi.org/10.1007/s10943-011-9481-9>.
38. Schmittdiel JA, Brown SD, Neugebauer R, et al. Health-plan and employer-based wellness programs to reduce diabetes risk: The Kaiser Permanente Northern California NEXT-D Study. *Prev Chronic Dis*. 2013; 10:E15. <http://dx.doi.org/10.5888/pcd10.120146>. [PubMed: 23369768]
39. Duru OK, Mangione CM, Chan C, et al. Evaluation of the diabetes health plan to improve diabetes care and prevention. *Prev Chronic Dis*. 2013; 10:E16. <http://dx.doi.org/10.5888/pcd10.120150>. [PubMed: 23369769]
40. Albu J, Sohler N, Matti-Orozco B, et al. Expansion of electronic health record-based screening, prevention, and management of diabetes in New York City. *Prev Chronic Dis*. 2013; 10:E13. <http://dx.doi.org/10.5888/pcd10.120148>. [PubMed: 23369766]
41. Kaiser Family Foundation. *Employer health benefits 2013 annual survey*. The Kaiser Family Foundation and Health Research & Educational Trust; Menlo Park, CA: 2013. <http://ehbs.kff.org/>
42. Haviland AM, Marquis MS, McDevitt RD, Sood N. Growth of consumer-directed health plans to one-half of all employer-sponsored insurance could save \$57 billion annually. *Health Aff (Millwood)*. 2012; 31(5):1009–1015. <http://dx.doi.org/10.1377/hlthaff.2011.0369>. [PubMed: 22566440]
43. Wharam JF, Galbraith AA, Kleinman KP, Soumerai SB, Ross-Degnan D, Landon BE. Cancer screening before and after switching to a high-deductible health plan. *Ann Intern Med*. 2008;

- 148(9):647–655. <http://dx.doi.org/10.7326/0003-4819-148-9-200805060-00004>. [PubMed: 18458277]
44. Wharam JF, Landon BE, Galbraith AA, Kleinman KP, Soumerai SB, Ross-Degnan D. Emergency department use and subsequent hospitalizations among members of a high-deductible health plan. *JAMA*. 2007; 297(10):1093–1102. <http://dx.doi.org/10.1001/jama.297.10.1093>. [PubMed: 17356030]
45. Wharam JF, Soumerai S, Trinacty C, et al. Impact of emerging health insurance arrangements on diabetes outcomes and disparities: rationale and study design. *Prev Chronic Dis*. 2013; 10:E11. <http://dx.doi.org/10.5888/pcd10.120147>. [PubMed: 23369764]
46. Ackermann RT, Holmes AM, Saha C. Designing a natural experiment to evaluate a national health care-community partnership to prevent type 2 diabetes. *Prev Chronic Dis*. 2013; 10:E12. <http://dx.doi.org/10.5888/pcd10.120149>. [PubMed: 23369765]




**Figure 1.** Relationships of NEXT-D Research Projects across the Landscape of Stakeholders that Influence Diabetes Prevention and Control.  
*Note:* The framework shown is an adaptation of socio-ecological models, depicting the spectrum, levels, and reach of different health system, societal, industry, and governmental stakeholder policies that may influence diabetes incidence, progression, and outcomes. The dashed lines separating stakeholders convey bidirectional and fluid interactions across levels. Numbered indicators represent the positioning of different NEXT-D studies, specifically:

1. Kaiser Permanente Northern California evaluation of health plan detection, outreach, and incentives for preventive care.
2. University of California Los Angeles evaluation of a disease specific health plan for patients with prediabetes or diabetes offered by employers to reduce costs of care.
3. New York St. Luke's Roosevelt Hospital Center evaluation of the use of an electronic medical record system for diabetes risk detection and prevention in community health centers and other outpatient clinics.
4. Harvard University evaluation of the effects on diabetes outcomes of employer-mandated switching to a high deductible health plan.
5. Northwestern University evaluation of the reach, adoption, implementation, effectiveness, and costs of a commercial health payer policy to support diabetes prevention programs delivered in communities.

Table 1

Hierarchy of Natural Experimental Study Designs and their Capacity to Address Threats to Internal Validity

Study design	Strength of design	Key strengths and weaknesses in ability to control for threats to validity*
Randomized controlled trial (experiment) Treatment: Randomize $O_1 \times O_2$ Comparison: Randomize $O_1 - O_2$	 <p>Strong</p> <p>Weak</p>	<ul style="list-style-type: none"> <li>● Gold standard of evidence</li> <li>● Controls for major validity threats* (history; maturation; selection; instrumentation)</li> <li>● Frequently not feasible to randomize individuals</li> <li>● Randomization of groups (e.g. clinics) is an option but can introduce bias; can lead to type 2 error (inadequate sample size); and requires statistical models for intercorrelation of data within groups</li> <li>● Often requires exclusion criteria that limits external generalizability</li> </ul>
Interrupted time series with a comparison series <sup>†</sup> Treatment: $O_1 O_2 O_3 \dots O_{12} X$ $O_{13} O_{14} O_{15} \dots O_{24}$ Comparison: $O_1 O_2 O_3 \dots O_{12} -$ $O_{13} O_{14} O_{15} \dots O_{24}$		<ul style="list-style-type: none"> <li>● Multiple baseline observations allows control for major validity threats* (history; maturation; selection; instrumentation)</li> </ul>
Single interrupted time series <sup>†</sup> Treatment: $O_1 O_2 O_3 \dots O_{12} X$ $O_{13} O_{14} O_{15} \dots O_{24}$		<ul style="list-style-type: none"> <li>● Baseline trends allow visible effects/control for most biases*</li> </ul>
Pre-post with comparison group Treatment: $O_1 \times O_2$ Comparison: $O_1 - O_2$		<ul style="list-style-type: none"> <li>● Comparability of baseline trend unknown</li> <li>● Strong risk of history, maturation, selection bias, regression to mean*</li> </ul>
Pre-post without comparison Treatment: $O_1 \times O_2$		<ul style="list-style-type: none"> <li>● Single observation before and after with no baseline trends</li> <li>● Weak design for assessing effectiveness</li> <li>● Strong risk of history, maturation, instrumentation bias, regression to mean*</li> </ul>
Cross-sectional designs Treatment: $X O_1$ Comparison: $O_1$		<ul style="list-style-type: none"> <li>● Weak design for assessing effectiveness</li> <li>● Simple correlation, no baseline, no measure of change</li> <li>● Very strong risk for most major threats to validity*</li> </ul>

Note: Readers are referred to Shadish, Cook, and Campbell (17) for a more extensive list of threats to validity.

X: represents intervention.

O: represents outcome measurement point.

\*Threats to Validity: *History*: an event occurring between pre- and post-tests when the event is not the intervention of research interest. *Maturation*: Subjects growing older, wiser, healthier (or sicker), etc. between pre- and post-test, & changes are not the intervention of interest. *Instrumentation*: a change in the measuring instrument (either humans or test). *Selection*: pre-intervention differences between people in one experimental group vs another; Selection biases can interact with maturation, history, and instrumentation biases, in which participants or groups selected as part of the intervention have different levels of maturation, history, or instrumentation than those in the control condition. *Statistical regression to the mean*: the tendency that when a measurement is extreme a second measurement will be closer to the mean.

<sup>†</sup>Time Series: In practice, a minimum of 12-24 observations (e.g., monthly rates) are usually required to measure discontinuities in outcomes while controlling for baseline trends (the counterfactual).